Unified Avionics Interconnect Standard Using Byte Addressable Scalable Coherent Interface/ Real Time

Open Systems Project Engineering Conference (OSPEC)

FY 98 Status Review

29 April - 1 May 1998

Ralph Lachenmaier
Naval Air Warfare Center/China Lake
ralph_lachenmaier@wssagw.chinalake.navy.mil
760-939-7832

Outline

- Byte Addressable Unified Avionics Interconnect Concept
- Why a Byte Addressable Shared Memory Type Network?
- Project Approach
- Performers–OS-JTF, SBIRs
- Some Sample Simulation Results from U of Florida
- Project Schedule
- Issues
- Summary

Byte Addressable/ Shared Memory Unified Avionics Interconnect Concept

GENERAL PROBLEM:

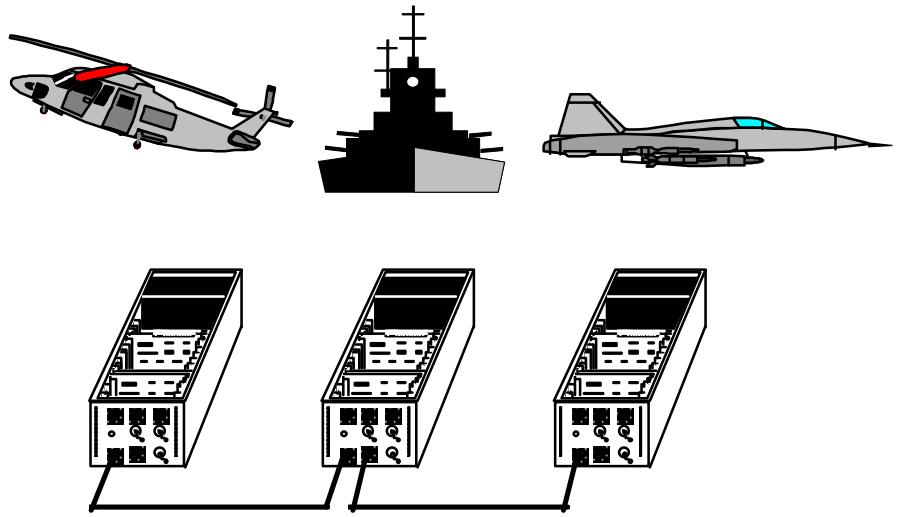
- Unifying avionics interconnects reduces costs
- Byte addressable networks unify backplanes and LANs while dramatically improving performance, but lack desirable real time features, e.g. Scalable Coherent Interface
- Limitations of current networks
  - Fibre Chan.,ATM: poor latency, not byte add.
  - Firewire: byte add., slow
  - SCI: byte add. fast, limited real time features

SOLUTION:

- SAE has selected SCI as base for Unified Avionics Interconnect Standard
- Work with SAE and IEEE standardization groups to establish a Byte Addressable Unified Avionics Interconnect Standard–SCI/Real Time
- Add features for determinism, priority, fault tolerance, and security

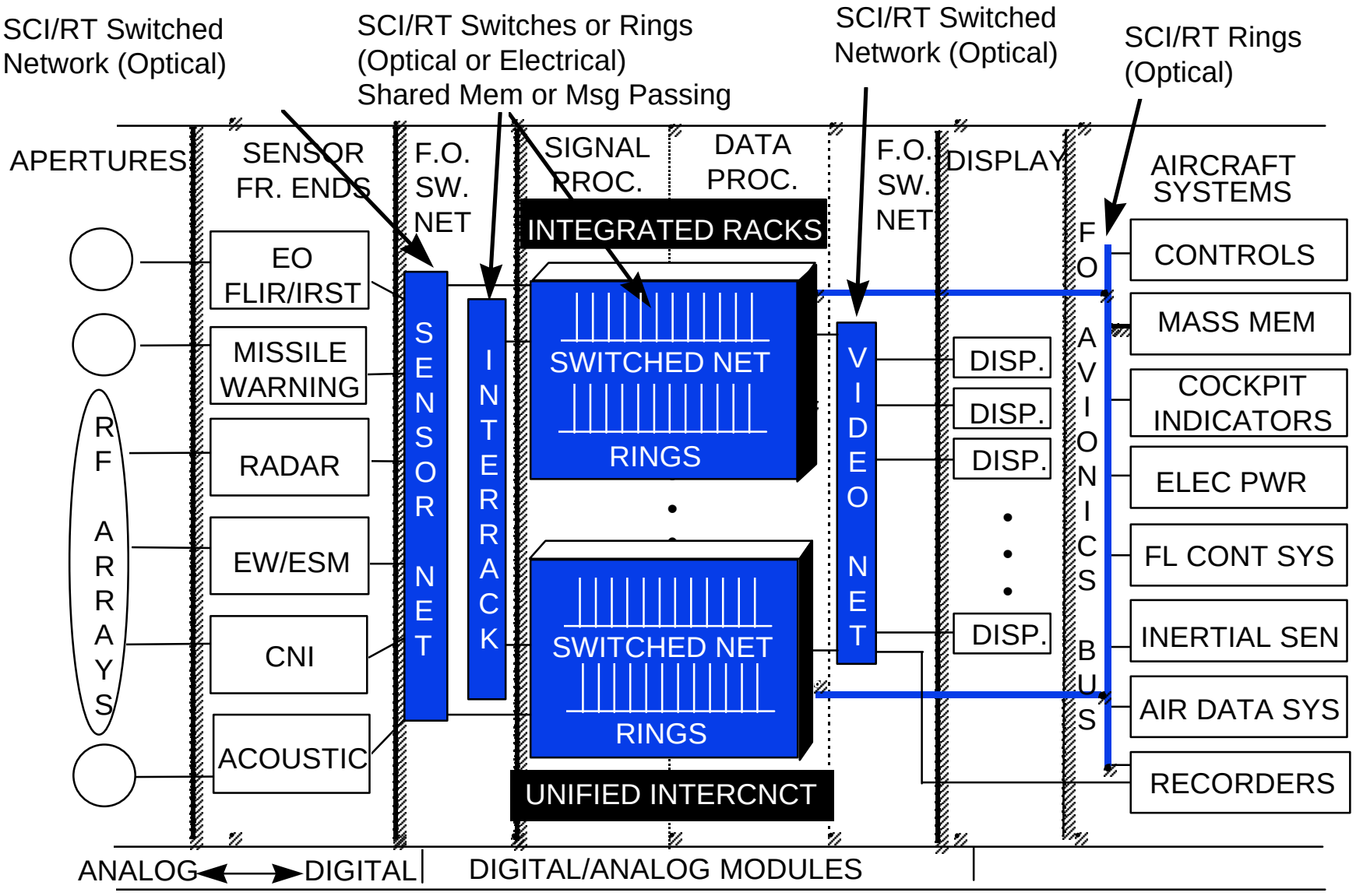
BENEFICIARIES:

- Any military real time program using an open architecture and requiring: high performance, low latency, unified interconnect.
  - Examples include: Joint Strike Fighter, UCAV, Integrated Sensor System, Canadian Navy, Norwegian Defense Research Establishment
- Commercial real time users such as telecom manufacturers



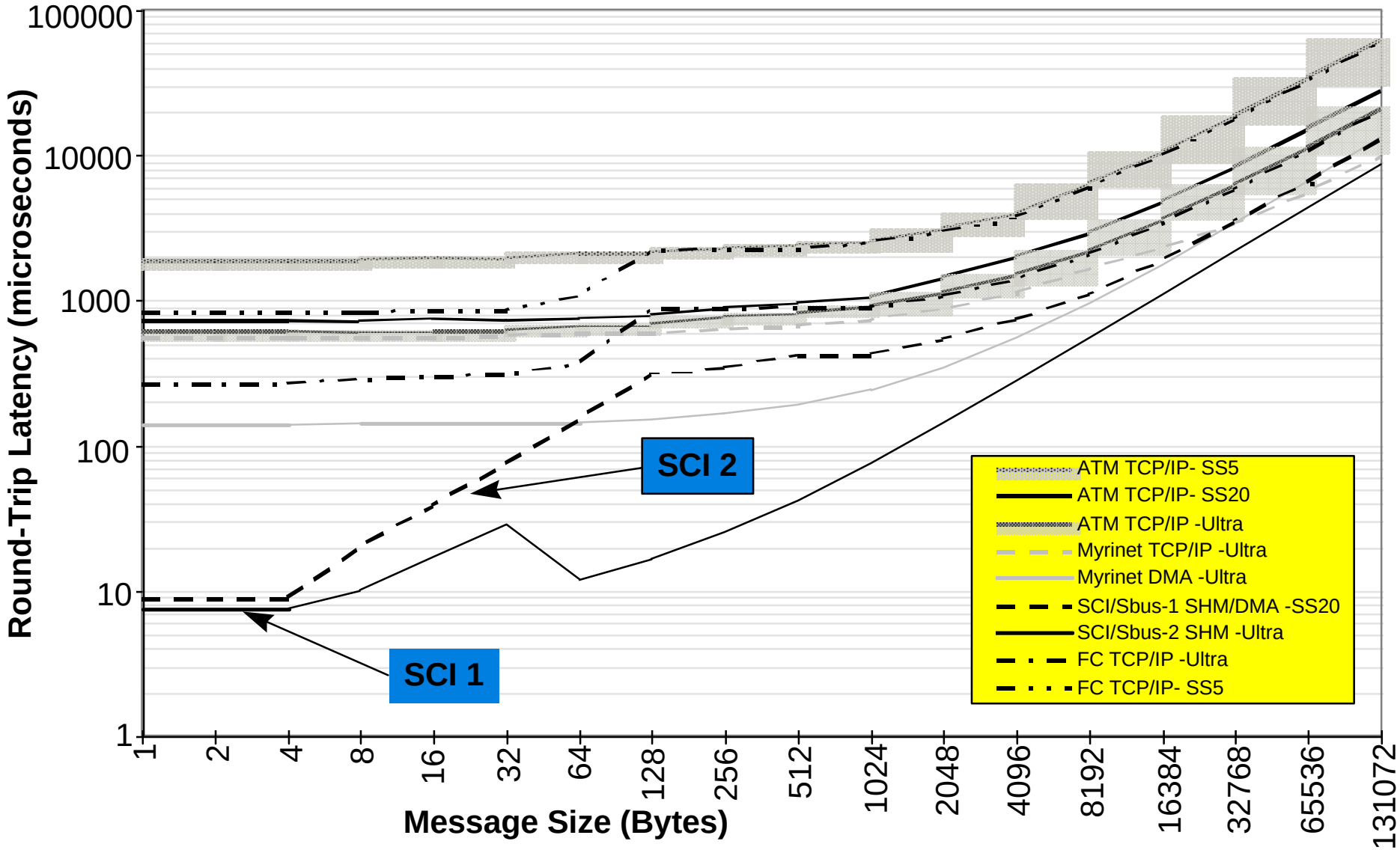
ARCHITECTURAL PERSPECTIVE

Open Architecture built on SCI Unified Network



Why a Byte Addressable Shared Memory Type Network?

Measured Latency of Different High-Performace Networks



Measured Performance Numbers of Note

Minimum Round Trip Latencies

- **SCI/Sbus-2 SHM Ultra:** **7.4 µs**
- **SCI/Sbus-1 SHM/DMA SS20:** **8.7 µs**
- **Myrinet DMA Ultra:** **139.3 µs**
- **FC TCP/IP Ultra:** **266.5 µs**
- **Myrinet TCP/IP Ultra:** **558.6 µs**
- **ATM TCP/IP Ultra:** **606.5 µs**
- **ATM TCP/IP SS20:** **713.7 µs**
- **FC TCP/IP SS5:** **840.1 µs**
- **ATM TCP/IP SS5:** **1919.1 µs**

Measurements were done on various Sun workstations thru the "S-BUS". The S-BUS is the limiting factor on further latency reductions. SCI latencies have been measured in the 1 µs range on other systems

Project Approach

OPEN SYSTEMS DEMONSTRATION

- Analysis and simulation of various SCI protocol enhancements to show proof of performance
- Results used by IEEE SCI/RT Working Group
- Analyzing and simulating various priority based flow control mechanisms
  - Analyzing various fault tolerance enhancements: Counter rotating rings with “ring wrap” on failure; Skip a node on failure; Duplicate suppression
  - With JSF, examined use of address field and page tables for security
- Develop and domonstrate a prototype SCI/RT protocol chip using SBIR funding

OPEN SYSTEMS STANDARDIZATION

- IEEE P1596.6 Scalable Coherent Interface for Real-Time Systems
  - Deterministic behavior (priorities)
  - Additional fault tolerance
  - Security mechanisms
  - As close to commercial as possible
- Work with SAE Avionics Systems Division

Performers

- **Two Main Performers**
  - **University of Florida-- High-Performance Computing and Simulation (HCS) Research Laboratory** (FY-96 $70K, FY-97 $50K)
    - Have a network lab supporting both experimental and simulation work
    - Have considerable experience with SCI; Have done performance studies and published papers
    - Emphasis on simulations, but also experimentation: BONeS, Networked Sun W/S
  - **Edgewater Computer Systems, Inc.** (FY-96 $80K)
    - Originated SCI/RT concept
    - Good background in Rate Monotonic Scheduling
    - Performed studies for the Canadian Navy
    - Wrote original draft standard
    - Emphasis will be on analysis with some simulation
    - Simulation system: Discrete Event Simulation (DEVS)
- **Support: Rich Fryer**—Writing standards documents (FY-97 $50K)

SBIR Contractors

- **Interconnect Systems Solutions**
  - Developing real time protocol chip for SCI
  - SBIR Phase I underway
  - Need support to ascertain Phase II
- **Daryoush GEMS**
  - Developing 5-10 Gbit/sec serial fiber optic link to be used on SCI or other interconnect
  - SBIR Phase I underway
  - Need support to ascertain Phase II
- **Combination of the two SBIRs would give outstanding performance:**
  - with very low module pin count
  - using same interconnect for backplane or rack to rack
  - supporting either shared memory or message passing
  - **using an IEEE standard interconnect**

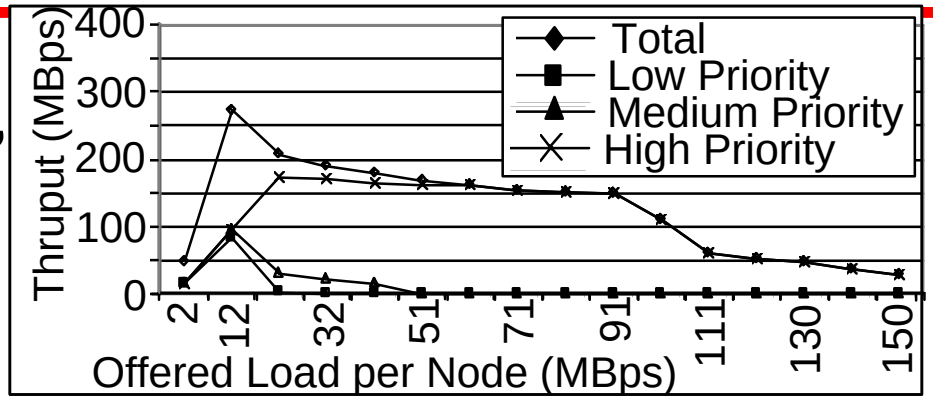
Some Sample Simulation Results from U of Florida

Candidate SCI/RT Protocols being Simulated

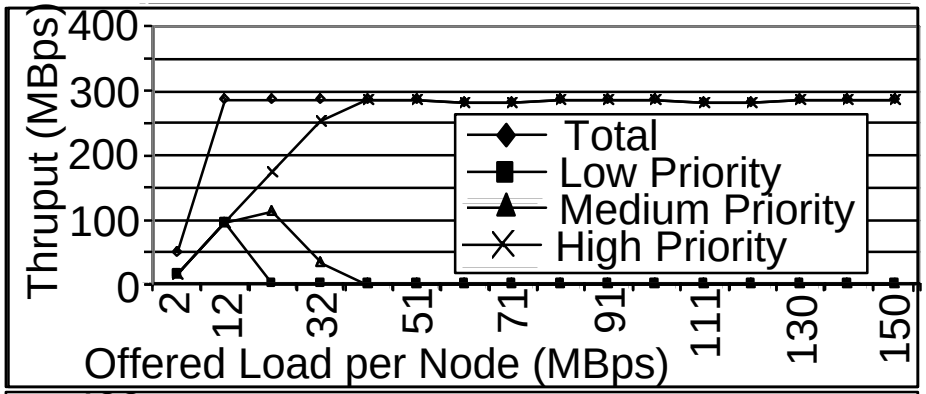
- **PPQ = Pre-emptive Priority Queue Protocol**
- **Train = Token-Train Protocol**
- **DFC = Directed Flow Control Protocol**

Multi-priority Exper. Results Thruput (Input Queue Stall Time = 1.0 us)

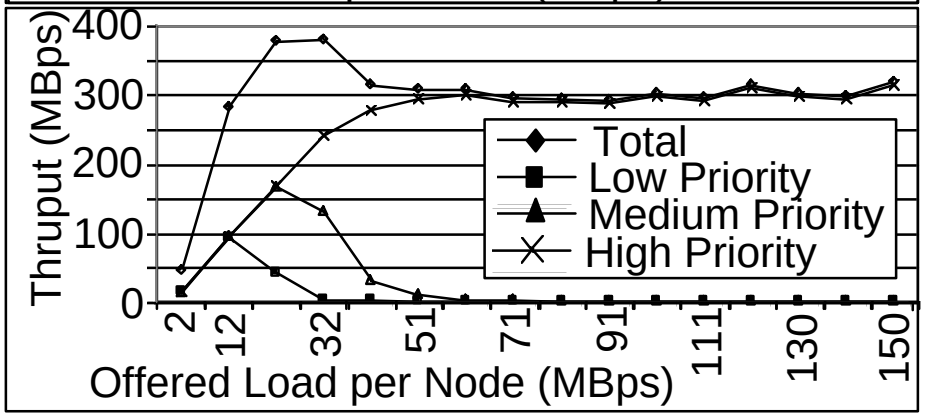
PPQ



TRAIN



DFC



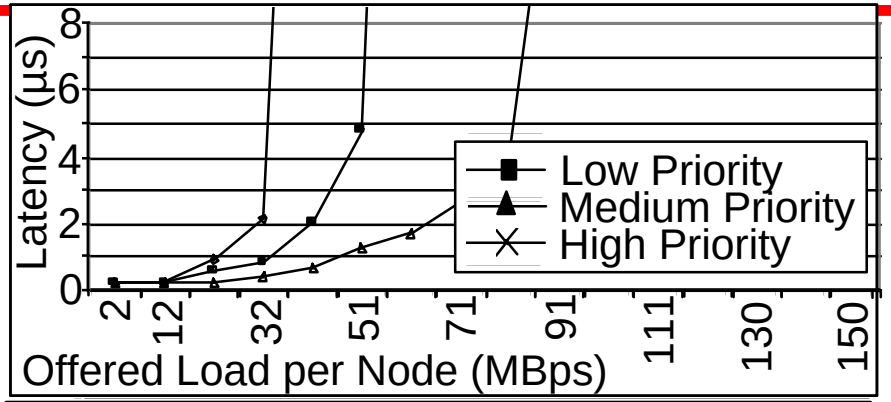
- With a slower IQST, the amount of bandwidth consumed by busy retries, primarily out of the input queue, overwhelms actual data transmission.

- Here, the primary limiting factor is the slow input queues not the arbitration mechanism. Thus, the TRAIN is able to achieve fair performance in comparison.

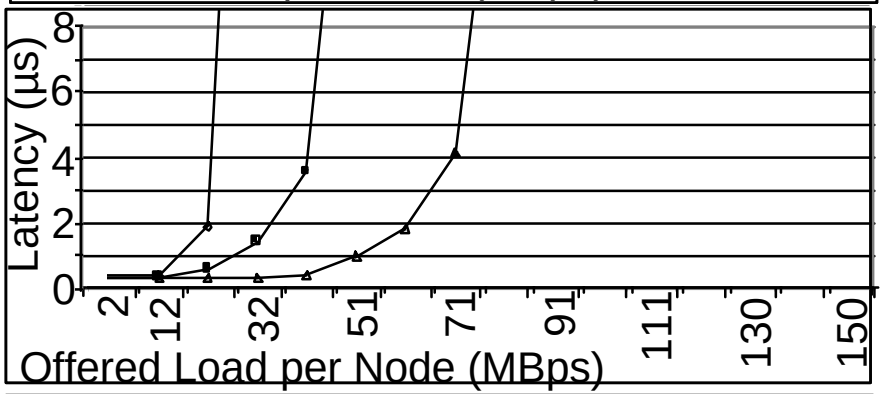
- Once again, the low overhead of DFC allows the highest sustained data throughput: about 300 MBps. The performance does not degrade as the loading increases as it does for PPQ.

Multi-priority Exper. Results Latency
(Input Queue Response Time = 0.1 us)

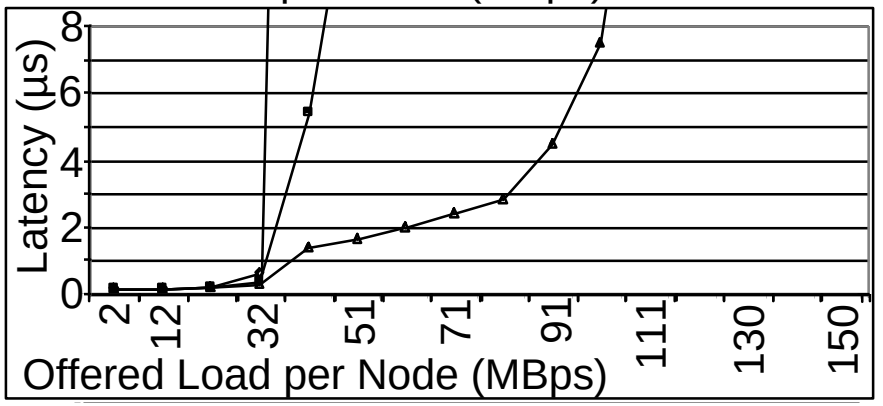
PPQ



TRAIN

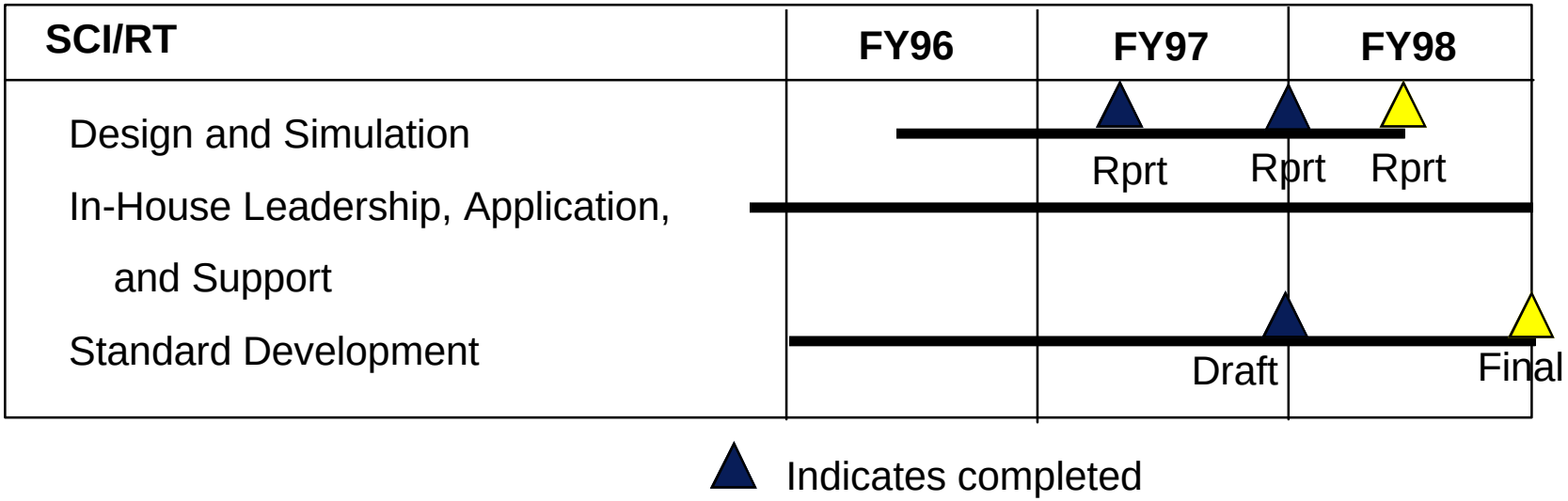


DFC



- PPQ performs well in this category having average one-way latencies of 0.20 us for all priority levels.
- The extra amount of overhead involved in transmission is evident in the TRAIN protocol's latency measurements. Average one-way latencies were 0.36 us for high-priority packets and slightly higher for low- and medium-priorities.
- DFC has the lowest minimum latency of all three protocols at 0.17 us for all three priority levels. The latency degradation is similar to that of PPQ, with the packets being held until there is room at the destination instead of being busy-retried.

Project Schedule
(Assuming FY-98 Funding)



This schedule has slipped because of no FY-98 funding to this time

Issues

- **No OS-JTF funding in received for FY-98 (only current activity is SBIRs)**
  - **U of FL simulation work left incomplete. More simulation needed to be assured of making best protocol selection**
  - **Work on IEEE 1596.6 Scalable Coherent Interface/ Real Time standard has stopped with standard in draft phase**
- **OS-JTF support sought for SCI/RT SBIRs**
  - **Two SCI/RT related SBIR ready for Phase II in June**
  - **OS-JTF endorsement would help guarantee transition**

Summary

- **A standard byte addressable interconnect has the potential to replace multiple present day avionics interconnects–providing a Unified Avionics Interconnect**
- **IEEE 1596 Scalable Coherent Interface (SCI) has the highest throughput and lowest latency of any of the commercial interconnects available today, and is the best starting point to meeting the requirements of a Unified Avionics Interconnect.**
- **IEEE 1596.6 SCI/RT adds determinism, priorities, security, and improved fault tolerance to base SCI.**
- **Further OS-JTF support is needed to complete SCI/RT**
  - **Financial support to complete simulations and write the standard**
  - **Programmatic support for SBIRs prototyping SCI/RT.**

BACKUPS

Byte addressable networks now catching on

- Firewire most popular: Microsoft, Intel, Sony have announced products
- SCI progressing more slowly: HP/Convex, Sequent, Data General, Siemens, Sun, Scali, (Cray, IBM: SCI like)
- Serial Express (SCI Derivative): Under development driven by Sun
- Sun S3mp research program uses proprietary byte addressable network

INTERCONNECT COMPARISON

Factor	SCI	Fibre Channel	ATM
Primary target usage	•Optimized for CPU to-memory and CPU-to-CPU interconnect within (board to board) or between cabinet(s)	•Optimized for host-to-peripheral and host-to-host channel usage	•Optimized for long haul tele-communications with spin offs for LAN application
Fundamental factors affecting usage	•Both node and byte addressable for direct memory access •Supports both message passing and shared memory (caching optional)	•Node addressable--does not support byte addresssing •Supports message passing	•Node addressable--does not support byte addressing •Supports message passing
Physical characteristics	•Serial or 16 bit parallel •Electrical or optical	•Serial •Electical or optical	•Serial •Electrical or optical
Bandwidth	•8 Gbps parallel or 1.25 Gbps serial •Higher speeds planned	•1.0 Gbps •Higher speeds planned	•622 Gbps •Higher speeds planned
Latency	•Current benchmark (Dolphin): Less than 5 micro-seconds between hosts Loral SMPS 5/96: About 1µs between hosts	•Current benchmark (U of FL) Approx. 200 micro-seconds between hosts	•Current benchmark (U of FL) Approx. 600 micro-seconds between hosts

INTERCONNECT COMPARISON(2)

Factor	SCI	Fibre Channel	ATM
Fundamental factors affecting performane	•Direct memory addressing bypasses software for 100 times latency improvement •Variable length packet up to 256 bytes of data •18 bytes overhead: 16-in packet, 2-b/w packets •Optimized for connectionless switching	•Node addressing requires cpu interrupts and software support •Variable length frame up to 2112 bytes of data •60 bytes overhead: 36-in frame, 24-b/w frames •Optimized for connection oriented switching	•Node addressing requires cpu interrupts and software support •Fixed length 53 byte packet--48 bytes data, 5 bytes header. •Additional framing packets required to transmit data •Virtual connection oriented switching
Flexibility	•Supports switched network •Supports insertion rings	•Supports switched network •Supports Arbitrated Loop topology	•Supports switched network •Ring support under development
Typical Users	•High performance servers and supercomputers: Hewlett Packard/ Convex, Sequent, Data General, Siemens, Scali, Cray. •Multi-workstation/ PC inter-connect	•Disk manufacturers and high performance LANs: SeaGate (and compatible hosts), Ancore, Systran, etc.	•Telephone companies (long haul and WAN): ATT, Ericksson, Alcatel, etc. High performance LANs: Fore, Telco, etc.